

# **OECD GUIDELINE FOR THE TESTING OF CHEMICALS**

## **DRAFT PROPOSAL FOR AN UPDATE TO TEST GUIDELINE 429**

### **Skin Sensitisation: Local Lymph Node Assay**

#### **INTRODUCTION**

1. OECD Guidelines for the Testing of Chemicals are periodically reviewed in the light of scientific progress, changing regulatory needs, and animal welfare considerations. The first Test Guideline (TG) for the determination of skin sensitisation in the mouse, the Local Lymph Node Assay (LLNA) TG (i.e., TG 429) was adopted in 2002 (1), after sufficient validation studies. The details of the validation of the LLNA and a review of the associated work have been published (2)(3)(4)(5)(6)(7)(8). The updated LLNA is based on the evaluation of experience and scientific data (9). This is the second Test Guideline to be promulgated for assessing skin sensitisation potential of chemicals in animals. The other Test Guideline (i.e., TG 406) utilises guinea pig tests, notably the guinea pig maximisation test and the Buehler test (10). This updated Test Guideline includes a set of Performance Standards (PS) (Annex 1) that can be used to more efficiently evaluate the validation status of new and/or modified test methods that are functionally and mechanistically similar to the LLNA, in accordance with the principles of Guidance Document No. 34 (11).

2. The LLNA studies the induction phase of skin sensitisation and provides quantitative data suitable for dose response assessment. It should be noted that the mild/moderate sensitisers, which are recommended as suitable positive control substances for guinea pig test methods, are also appropriate for use with the LLNA (6)(8)(12). The LLNA provides certain advantages with regard to animal welfare and a reduced LLNA (rLLNA) for hazard classification of skin sensitising substances can be performed under this Test Guideline (13)(14).

#### **DEFINITIONS**

3. Definitions used are provided in Annex 2.

#### **INITIAL CONSIDERATIONS**

4. The LLNA provides an alternative method for identifying skin sensitising chemicals and for confirming that chemicals lack a significant potential to cause skin sensitisation. This does not necessarily imply that in all instances the LLNA should be used in place of guinea pig tests, but rather that the assay is of equal merit and may be employed as an alternative in which positive and negative results generally no longer require further confirmation.

5. The LLNA is an *in vivo* method and, as a consequence, will not eliminate the use of animals in the assessment of allergic contact sensitising activity. It has, however, the potential to reduce the number of animals required for this purpose. Moreover, the LLNA offers a substantial refinement of the way in which animals are used for allergic contact sensitisation testing. The LLNA is based upon consideration of immunological events

stimulated by chemicals during the induction phase of sensitisation. Unlike guinea pig tests (i.e., TG 406) (10) the LLNA does not require that challenge-induced dermal hypersensitivity reactions be elicited. Furthermore, the LLNA does not require the use of an adjuvant, as is the case for the guinea pig maximisation test, as described in reference (10). Thus, the LLNA reduces animal distress. Despite the advantages of the LLNA over TG 406, it should be recognised that there are certain limitations that may necessitate the use of TG 406 (e.g., false negative findings in the LLNA with certain metals, false positive findings with certain skin irritants) (15).

## **PRINCIPLE OF THE TEST**

6. The basic principle underlying the LLNA is that sensitisers induce proliferation of lymphocytes in the lymph nodes draining the site of chemical application. This proliferation is proportional to the dose and to the potency of the applied allergen and provides a simple means of obtaining a quantitative measurement of sensitisation. The LLNA assesses this proliferation as the proliferation in test groups compared to that in vehicle treated controls. The ratio of the proliferation in treated groups to that in the concurrent vehicle control group, termed the Stimulation Index (SI), is determined, and should be  $\geq 3$  before a test substance can be further evaluated as a potential skin sensitizer. The methods described here are based on the use of *in vivo* radioactive labelling to measure an increased number of proliferating cells in the draining auricular lymph nodes. However, other endpoints for assessment of the number of proliferating cells may be employed provided there is justification and appropriate scientific support, including full citations and description of the methodology.

## **DESCRIPTION OF THE ASSAY**

### **Selection of animal species**

7. The mouse is the species of choice for this test. Young adult female mice of CBA/Ca or CBA/J strain, which are nulliparous and non-pregnant, are used. At the start of the study, animals should be between 8-12 weeks old, and the weight variation of the animals should be minimal and not exceed 20% of the mean weight. Other strains and males may be used when sufficient data are generated to demonstrate that significant strain and/or gender-specific differences in the LLNA response do not exist.

### **Housing and feeding conditions**

8. Mice should be group housed (16), unless adequate scientific rationale for housing mice individually is provided. The temperature of the experimental animal room should be 22°C ( $\pm$  3°C). Although the relative humidity should be at least 30% and preferably not exceed 70%, other than during room cleaning, the aim should be 50-60%. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water.

### **Preparation of animals**

9. The animals are randomly selected, marked to permit individual identification (but not by any form of ear marking), and kept in their cages for at least five days prior to the start

of dosing to allow for acclimatisation to the laboratory conditions. Prior to the start of treatment all animals are examined to ensure that they have no observable skin lesions.

#### **Preparation of dosing solutions**

10. Solid test substances should be dissolved in appropriate solvents/vehicles and diluted, if appropriate, prior to application to an ear of the mice. Liquid test substances may be applied neat or diluted prior to dosing. Insoluble materials, such as those generally seen in medical devices, should be extracted in an appropriate solvent and, if appropriate, further processed prior to application to an ear of the mice. The test substance should be prepared daily unless stability data demonstrate the acceptability of storage.

#### **Reliability check**

11. Positive controls are used to demonstrate appropriate performance of the assay by responding with adequate and reproducible sensitivity to a sensitising substance for which the magnitude of the response is well characterised. Inclusion of a concurrent positive control is recommended because it demonstrates competency of the laboratory to successfully conduct each assay and allows for an assessment of intra- and inter-laboratory reproducibility and comparability. The positive control should produce a positive LLNA response at an exposure level expected to give an increase in the SI > 3 over the negative control group. The positive control dose should be chosen such that the induction is reproducible but not excessive. Preferred positive control substances are hexyl cinnamic aldehyde (Chemical Abstracts Service [CAS] No 101-86-0) and mercaptobenzothiazole (CAS No 149-30-4). There may be circumstances in which, given adequate justification, other positive control substances, meeting the above criteria, may be used.

12. While inclusion of a concurrent positive control group is recommended, there may be situations in which periodic testing (i.e., at intervals  $\leq 6$  months) of the positive control substance may be adequate for laboratories that conduct the LLNA regularly (i.e., conduct the LLNA at a frequency of no less than once per month) and have an established historical positive control database that demonstrates the laboratory's ability to obtain reproducible and accurate results with positive controls. Adequate proficiency with the LLNA can be successfully demonstrated by generating consistent results with the positive control in at least 10 independent tests conducted within a reasonable period of time (i.e., less than one year).

13. A concurrent positive control group should always be included when there is a procedural change to the LLNA (e.g., change in trained personnel, change in test method materials and/or reagents, change in test method equipment, change in source of test animals), and such changes should be documented in laboratory reports. Consideration should be given to the impact of these changes on the adequacy of the previously established historical database in determining the necessity for establishing a new historical database to document consistency in the positive control results.

14. Investigators should be aware that the decision to conduct a positive control on a periodic basis instead of concurrently has ramifications on the adequacy and acceptability of negative study results generated without a concurrent positive control during the interval

between each periodic positive control study. For example, if a false negative result is obtained in the periodic positive control study, all negative test substance results obtained in the interval between the last acceptable periodic positive control study and the unacceptable periodic positive control study will be questioned. Any study reports associated with these negative test substance results should immediately be amended to report the failed positive control test. In order to demonstrate that the prior negative test substance study results are acceptable, a laboratory would be expected to repeat all negative studies, which would require additional expense and increased animal use. Simply repeating a failed periodic positive control study is not scientifically valid. These implications should be carefully considered when determining whether to include concurrent positive controls or to only conduct periodic positive controls. Consideration should also be given to using fewer animals in the concurrent positive control group when this is scientifically justified and if the laboratory demonstrates, based on laboratory-specific historical data, that fewer mice can be used without substantially increasing the frequency with which studies will need to be repeated.

15. Although the positive control substance should be tested in the vehicle that is known to elicit a consistent response (e.g., acetone: olive oil), there may be certain regulatory situations in which testing in a non-standard vehicle (clinically/chemically relevant formulation) will also be necessary. In such situations the possible interaction of a positive control with this unconventional vehicle should be tested. If the concurrent positive control substance is tested in a different vehicle than the test substance, then a separate vehicle control for the concurrent positive control should be included.

16. In instances where substances of a specific chemical class or range of responses are being evaluated, benchmark controls may be useful to demonstrate that the test method is functioning properly for detecting the skin sensitisation potential of a test substance. Appropriate benchmark controls should have the following properties:

- structural and functional similarity to the class of the substance being tested;
- known physical/chemical characteristics;
- supporting data on known effects in animal models;
- known potency for sensitisation response.

## **TEST PROCEDURE**

### **Number of animals and dose levels**

17. A minimum of four animals is used per dose group, with a minimum of three concentrations of the test substance, plus a concurrent negative control group treated only with the vehicle for the test substance, and a concurrent positive control (see paragraphs 11-15). Except for absence of treatment with the test substance, animals in the control groups should be handled and treated in a manner identical to that of animals in the treatment groups.

18. Dose and vehicle selection should be based on the recommendations given in references (3) and (5). Doses are selected from the concentration series 100%, 50%, 25%, 10%, 5%, 2.5%, 1%, 0.5%, etc. Existing acute toxicity and dermal irritation data should be considered, where available, in selecting the three consecutive concentrations so that the highest concentration maximises exposure whilst avoiding systemic toxicity and excessive local skin irritation (3)(17). In the absence of such information, an initial prescreen test may be necessary (see paragraphs 21-24).

19. The vehicle should not interfere with or bias the test result and should be selected on the basis of maximising the solubility in order to obtain the highest concentration achievable whilst producing a solution/suspension suitable for application of the test substance. In order of preference, recommended vehicles are acetone: olive oil (4:1 v/v), *N,N*-dimethylformamide, methyl ethyl ketone, propylene glycol, and dimethyl sulphoxide (3)(8), but others may be used if sufficient scientific rationale is provided. In certain situations it may be necessary to use a clinically relevant solvent or the commercial formulation in which the test substance is marketed as an additional control. Particular care should be taken to ensure that hydrophilic materials are incorporated into a vehicle system, which wets the skin and does not immediately run off. Thus, wholly aqueous vehicles are to be avoided.

20. The processing of lymph nodes from individual mice allows for the assessment of interanimal variability and a statistical comparison of the difference between test substance and vehicle control group measurements. In addition, evaluating the possibility of reducing the number of mice in the positive control group is only feasible when individual animal data are collected.

#### **Prescreen test**

21. The purpose of the prescreen test is to provide guidance for selecting the maximum dose level to use in the main LLNA study. The maximum dose level tested should be a concentration of 100% (i.e., neat substance) for liquid substances or the maximum soluble concentration (for solids), unless available information suggests that this concentration induces systemic toxicity or excessive local irritation after topical application in the mouse.

22. In the absence of such information, a prescreen test should be performed using three dose levels of the test substance, in order to define the appropriate dose level to test in the LLNA. Six mice (two per concentration) are used, and the prescreen test is conducted under identical conditions as the main LLNA study, except there is no assessment of lymph node proliferation. All mice will be observed daily for any clinical signs of systemic toxicity or local irritation at the application site. Body weights are recorded pre-test and prior to termination (Day 6). Both ears of each mouse are observed for erythema and scored using Table 1. Ear thickness measurements are taken using a thickness gauge (e.g., digital micrometer or Peacock Dial thickness gauge) on Day 1 (pre-dose), Day 3 (approximately 48 hours after the first dose), and Day 6. Excessive local irritation is indicated by an erythema score  $\geq 3$  and/or ear swelling of  $\geq 25\%$  (18)(19).

194 **Table 1 Erythema Scores**

Observation	Value
No visual effect	0
Slight erythema (barely perceptible)	1
Well-defined erythema	2
Moderate to severe erythema (beet redness)	3
Eschar (i.e., piece of dead tissue that is cast off from the surface of the skin)	4

195 23. In addition to a 25% increase in ear swelling (18)(19), a statistically significant  
196 increase in ear swelling in the treated mice compared to control mice has also been used to  
197 identify irritants in the LLNA (20)(21)(22)(23)(24)(25)(26). While statistically significant  
198 increases can occur when ear swelling is less than 25%, they have not been associated  
199 specifically with excessive irritation (22)(24)(25)(26). Additionally, an adequately robust  
200 statistical comparison would require that a vehicle control group be included and that more  
201 than two mice per group be tested. Both of these requirements would substantially increase  
202 the number of mice used in a prescreen test. For this reason, a threshold increase in ear  
203 swelling above pre-dosing levels is recommended for this prescreen test.

204 24. Test guidelines for assessing acute dermal toxicity recommend a number of clinical  
205 observations for assessing systemic toxicity (27)(28). The following clinical observations,  
206 which are based on test guidelines and current practices (29), may indicate systemic toxicity  
207 when used as part of an integrated assessment and therefore may indicate the maximum dose  
208 level to use in the main LLNA:

- 209 • Changes in nervous system function (e.g., piloerection, ataxia, tremors, and  
210 convulsions)
- 211 • Changes in behavior (e.g., aggressiveness, change in grooming activity,  
212 marked change in activity level)
- 213 • Changes in respiratory patterns (i.e., changes in frequency and intensity of  
214 breathing such as dyspnea, gasping, and rales)
- 215 • Changes in food and water consumption
- 216 • Lethargy and/or unresponsiveness
- 217 • Any clinical signs of more than slight or momentary pain and distress
- 218 • Reduction in body weight >10% from Day 1 to Day 6
- 219 • Mortality

220 **Reduced LLNA**

25. Use of a rLLNA protocol (13)(14) has the potential to reduce the number of animals used in the LLNA by omitting the middle and low dose groups. This is the only difference between the LLNA and the rLLNA and thus, the test substance concentration evaluated in the rLLNA should be the maximum concentration that does not induce overt systemic toxicity and/or excessive local irritation in the mouse. The rLLNA should be used for the hazard classification of skin sensitising substances if dose-response information is not needed, provided there is adherence to all other LLNA protocol specifications, as described in this Test Guideline. To further reduce animal use, the rLLNA should be used routinely as an initial test to determine allergic contact dermatitis potential of chemicals and products before conducting the LLNA. Negative substances can be classified as nonsensitisers and positive substances can be classified as sensitisers.

### **Main study experimental schedule**

26. The experimental schedule of the assay is as follows:

- *Day 1:*  
Individually identify and record the weight of each animal and any clinical observations. Apply 25 µL of the appropriate dilution of the test substance, the vehicle alone, or the concurrent positive control (see paragraphs 11-15), to the dorsum of each ear.
- *Days 2 and 3:*  
Repeat the application procedure carried out on Day 1.
- *Days 4 and 5:*  
No treatment.
- *Day 6:*  
Record the weight of each animal. Inject 250 µL of sterile phosphate-buffered saline (PBS) containing 20 µCi (7.4e+5 Bq) of tritiated (<sup>3</sup>H)-methyl thymidine into all test and control mice via the tail vein. Alternatively, inject 250 µL sterile PBS containing 2 µCi (7.4e + 4 Bq) of <sup>125</sup>I-iododeoxyuridine and 10<sup>-5</sup>M fluorodeoxyuridine into all mice via the tail vein. Five hours (5 h) later, humanely kill the animals. Excise the draining auricular lymph nodes from each mouse ear and process separately in PBS for each animal. Details and diagrams of the node identification and dissection can be found in reference (9).

### **Preparation of cell suspensions**

27. A single cell suspension of lymph node cells (LNC) excised bilaterally from each mouse is prepared by gentle mechanical disaggregation through 200 micron-mesh stainless steel gauze or another acceptable technique for generating a single-cell suspension. LNC are washed twice with an excess of PBS and the DNA is precipitated with 5% trichloroacetic acid (TCA) at 4°C for 18h (3). Pellets are either re-suspended in 1 mL TCA and transferred to scintillation vials containing 1.0 mL of scintillation fluid for <sup>3</sup>H-counting, or transferred directly to gamma counting tubes for <sup>125</sup>I-counting.

### **Determination of cellular proliferation (incorporated radioactivity)**

28. Incorporation of  $^3\text{H}$ -methyl thymidine is measured by  $\beta$ -scintillation counting as disintegrations per minute (DPM). Incorporation of  $^{125}\text{I}$ -iododeoxyuridine is measured by  $^{125}\text{I}$ -counting and also is expressed as DPM. The incorporation is expressed as DPM/mouse.

## **OBSERVATIONS**

### **Clinical observations**

29. Each mouse should be carefully observed once daily for any clinical signs, either of local irritation at the application site or of systemic toxicity. All observations are systematically recorded with records being maintained for each mouse. Monitoring plans should include criteria to promptly identify those mice exhibiting systemic toxicity, excessive irritation, or corrosion of skin for euthanasia.

### **Body weights**

30. As stated in paragraph 26, individual animal body weights should be measured at the start of the test and at the scheduled kill.

## **CALCULATION OF RESULTS**

31. Results for each treatment group are expressed as the mean SI. The SI is derived by dividing the mean DPM/mouse within each test substance group and the concurrent positive control group by the mean DPM/mouse for the solvent/vehicle control group. The average SI for vehicle treated controls is then one.

32. Collecting radioactivity data at the level of the individual mouse will enable a statistical analysis for presence and degree of dose response in the data. Any statistical assessment should include an evaluation of the dose response relationship as well as suitably adjusted comparisons of test groups (e.g., pair-wise dosed group versus concurrent vehicle control comparisons). Statistical analyses may include, for instance, linear regression or William's test to assess dose-response trends, and Dunnett's test for pairwise comparisons. In choosing an appropriate method of statistical analysis, the investigator should maintain an awareness of possible inequalities of variances and other related problems that may necessitate a data transformation or a non-parametric statistical analysis. In any case, the investigator should be alert to possible "outlier" responses for individual mice within a group that may necessitate analysis both with and without outliers.

33. The decision process with regard to a positive response includes a  $\text{SI} \geq 3$ , together with consideration of dose response and, where appropriate, statistical significance (4)(5)(6)(28).

34. If it is necessary to clarify the results obtained, consideration should be given to various properties of the test substance, including whether it has a structural relationship to known skin sensitizers, whether it causes excessive skin irritation in the mouse, and the nature of the dose response seen. These and other considerations are discussed in detail elsewhere (7).

299 **DATA AND REPORTING**

300 **Data**

301 35. Data should be summarised in tabular form showing the individual animal DPM  
302 values, the group mean DPM/animal, its associated error term, and the mean SI for each dose  
303 group compared against the concurrent vehicle control group.

304 **Test report**

305 36. The test report should contain the following information:

306 Test substance and control substances:

- 307 – identification data (e.g. CAS number, if available; source; purity; known  
308 impurities; lot number);
- 309 – physical nature and physicochemical properties (e.g. volatility, stability,  
310 solubility);
- 311 – if mixture, composition and relative percentages of components.

312 Solvent/vehicle:

- 313 – identification data (purity; concentration, where appropriate; volume used);
- 314 – justification for choice of vehicle.

315 Test animals:

- 316 – source of CBA mice;
- 317 – microbiological status of the animals, when known;
- 318 – number and age of animals;
- 319 – source of animals, housing conditions, diet, etc.

320 Test conditions:

- 321 – details of test substance preparation and application;
- 322 – justification for dose selection (including results from range finding study, if  
323 conducted);
- 324 – vehicle and test substance concentrations used, and total amount of substance  
325 applied;
- 326 – details of food and water quality (including diet type/source, water source);
- 327 – details of treatment and sampling schedules;
- 328 – methods for measurement of toxicity;
- 329 – criteria for considering studies as positive or negative;
- 330 – details of any protocol deviations and an explanation on how the deviation  
331 affects the study design and results.

332 Reliability check:

- 333 – a summary of results of latest reliability check, including information on
- 334 substance, concentration and vehicle used;
- 335 – concurrent and/or historical positive and negative control data for testing
- 336 laboratory;
- 337 – if a concurrent positive control was not included, the date and laboratory
- 338 report for the most recent periodic positive control and a report detailing the
- 339 historical positive control data for the laboratory justifying the basis for not
- 340 conducting a concurrent positive control.

341 Results:

- 342 – individual weights of mice at start of dosing and at scheduled kill; as well as
- 343 mean and associated error term for each treatment group.
- 344 – time course of onset and signs of toxicity, including dermal irritation at site of
- 345 administration, if any, for each animal;
- 346 – a table of individual mouse DPM values and SIs for each treatment group;
- 347 – mean and associated error term for DPM/mouse for each treatment group and
- 348 the results of outlier analysis for each treatment group;
- 349 – calculated SI and an appropriate measure of variability that takes into account
- 350 the interanimal variability in both the test substance and control groups;
- 351 – dose response relationship;
- 352 – statistical analysis, where appropriate.

353 Discussion of results:

- 354 – a brief commentary on the results, the dose-response analysis, and statistical
- 355 analyses, where appropriate, with a conclusion as to whether the test
- 356 substance should be considered a skin sensitiser.

357 Quality assurance statement for Good Laboratory Practice compliant studies:

- 358 – statement should indicate all inspections made during the study and the dates
- 359 any results were reported to the Study Director. The statement should also
- 360 confirm that the final report reflects the raw data.

361

**LITERATURE**

- (1) OECD (2002). Test Guideline 429. OECD Guideline for the Testing of Chemicals. Skin Sensitisation: Local Lymph Node Assay. Available at: [http://www.oecd.org/document/40/0,2340,en\_2649\_34377\_37051368\_1\_1\_1\_1,00.html].
- (2) Kimber, I. and Basketter, D.A. (1992). The murine local lymph node assay; collaborative studies and new directions: A commentary. Food and Chemical Toxicology, 30, 165-169.
- (3) Kimber, I., Dearman, R.J., Scholes, E.W. and Basketter, D.A. (1994). The local lymph node assay: developments and applications. Toxicology, 93, 13-31.
- (4) Kimber, I., Hilton, J., Dearman, R.J., Gerberick, G.F., Ryan, C.A., Basketter, D.A., Lea, L., House, R.V., Ladies, G.S., Loveless, S.E. and Hastings, K.L. (1998). Assessment of the skin sensitisation potential of topical medicaments using the local lymph node assay: An interlaboratory exercise. Journal of Toxicology and Environmental Health, 53, 563-79.
- (5) Chamberlain, M. and Basketter, D.A. (1996). The local lymph node assay: status of validation. Food and Chemical Toxicology, 34, 999-1002.
- (6) Basketter, D.A., Gerberick, G.F., Kimber, I. and Loveless, S.E. (1996). The local lymph node assay: A viable alternative to currently accepted skin sensitisation tests. Food and Chemical Toxicology, 34, 985-997.
- (7) Basketter, D.A., Gerberick, G.F. and Kimber, I. (1998). Strategies for identifying false positive responses in predictive sensitisation tests. Food and Chemical Toxicology, 36, 327-33.
- (8) Van Och, F.M.M., Slob, W., De Jong, W.H., Vandebriel, R.J. and Van Loveren, H. (2000). A quantitative method for assessing the sensitising potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of uncertainty margins. Toxicology, 146, 49-59.
- (9) ICCVAM (2009). Recommended Performance Standards: Murine Local Lymph Node Assay. NIH Publication Number 09-7357. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [http://iccvam.niehs.nih.gov/docs/immunotox\_docs/llna-ps/LLNAPerfStds.pdf].
- (10) OECD (1992). Test Guideline 406. OECD Guideline for Testing of Chemicals. Skin Sensitisation. Available at: [http://www.oecd.org/document/40/0,2340,en\_2649\_34377\_37051368\_1\_1\_1\_1,00.html].

- (11) OECD (2005). OECD Series on Testing and Assessment No. 34. Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Available at: [\[http://www.oecd.org/document/30/0,3343,en\\_2649\\_34377\\_1916638\\_1\\_1\\_1\\_1,00.html\]](http://www.oecd.org/document/30/0,3343,en_2649_34377_1916638_1_1_1_1,00.html).
- (12) Dearman, R.J., Hilton, J., Evans, P., Harvey, P., Basketter, D.A. and Kimber, I. (1998). Temporal stability of local lymph node assay responses to hexyl cinnamic aldehyde. *Journal of Applied Toxicology*, 18, 281-284.
- (13) Kimber, I., Dearman, R.J., Betts, C.J., Gerberick, G.F., Ryan, C.A., Kern, P.S., Patlewicz, G.Y. and Basketter, D.A. (2006). The local lymph node assay and skin sensitization: a cut-down screen to reduce animal requirements? *Contact Dermatitis*, 54, 181-185.
- (14) ESAC (2007). Statement on the Reduced Local Lymph Node Assay (rLLNA). European Commission Directorate General, Joint Research Centre, Institute for Health and Consumer Protection, European Centre for the Validation of Alternative Methods. April 2007. Available at: [\[http://ecvam.jrc.it/ft\\_doc/ESAC26\\_statement\\_rLLNA\\_20070525-1.pdf\]](http://ecvam.jrc.it/ft_doc/ESAC26_statement_rLLNA_20070525-1.pdf).
- (15) ICCVAM (1999). The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds. The Results of an Independent Peer Review Evaluation Coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No. 99-4494. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [\[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/llna/llnarep.pdf\]](http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf).
- (16) Institute of Laboratory Animal Research (ILAR) (1996). Guide for the Care and Use of Laboratory Animals. 7th ed. Washington, DC: National Academies Press.
- (17) OECD (2002). Test Guideline 404. OECD Guideline for Testing of Chemicals: Acute Dermal Irritation/Corrosion. Available at: [\[http://www.oecd.org/document/40/0,3343,en\\_2649\\_34377\\_37051368\\_1\\_1\\_1\\_1,00.html\]](http://www.oecd.org/document/40/0,3343,en_2649_34377_37051368_1_1_1_1,00.html)
- (18) Reeder, M.K., Broomhead, Y.L., DiDonato, L. and DeGeorge, G.L. (2007). Use of an enhanced local lymph node assay to correctly classify irritants and false positive substances. *Toxicologist*, 96 (S-1), 235.
- (19) ICCVAM (2009). Non-radioactive Murine Local Lymph Node Assay: Flow Cytometry Test Method Protocol (LLNA: BrdU-FC) Revised Draft Background Review Document. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [\[http://iccvam.niehs.nih.gov/methods/immunotox/fcLLNA/BRDcomplete.pdf\]](http://iccvam.niehs.nih.gov/methods/immunotox/fcLLNA/BRDcomplete.pdf).

- 436 (20) Hayes, B.B., Gerber, P.C., Griffey, S.S. and Meade, B.J. (1998). Contact  
437 hypersensitivity to dicyclohexylcarbodiimide and diisopropylcarbodiimide in female  
438 B6C3F1 mice. *Drug Chem Toxicol*, 21, 195-206.
- 439 (21) Homey, B., von Schilling, C., Blumel, J., Schuppe, H.C., Ruzicka, T., Ahr, H.J.,  
440 Lehmann, P. and Vohr, V.W. (1998). An integrated model for the differentiation of  
441 chemical-induced allergic and irritant skin reactions. *Toxicol Appl Pharmacol*, 153,  
442 83-94.
- 443 (22) Woolhiser, M.R., Hayes, B.B. and Meade, B.J. (1998). A combined murine local  
444 lymph node and irritancy assay to predict sensitization and irritancy potential of  
445 chemicals. *Toxicol Meth*, 8, 245-256.
- 446 (23) Hayes, B.B. and Meade, B.J. (1999). Contact sensitivity to selected acrylate  
447 compounds in B6C3F1 mice: relative potency, cross reactivity, and comparison of  
448 test methods. *Drug Chem Toxicol*, 22, 491-506.
- 449 (24) Ehling, G., Hecht, M., Heusener, A., Huesler, J., Gamer, A.O., van Loveren, H.,  
450 Maurer, T., Riecke, K., Ullmann, L., Ulrich, P., Vandebriel, R. and Vohr, H.W.  
451 (2005). A European inter-laboratory validation of alternative endpoints of the  
452 murine local lymph node assay: first round. *Toxicology*, 212, 60-68.
- 453 (25) Vohr, H.W. and Ahr, H.J. (2005). The local lymph node assay being too sensitive?  
454 *Arch Toxicol*, 79, 721-728.
- 455 (26) Patterson, R.M., Noga, E. and Germolec, D. (2007). Lack of evidence for contact  
456 sensitization by *Pfisteria* extract. *Environ Health Perspect*, 115, 1023-1028.
- 457 (27) OECD (1987). Test Guideline 402. OECD Guideline for Testing of Chemicals  
458 Acute Dermal Toxicity. Available at:  
459 [[http://www.oecd.org/document/40/0,3343,en\\_2649\\_34377\\_37051368\\_1\\_1\\_1\\_1,00.](http://www.oecd.org/document/40/0,3343,en_2649_34377_37051368_1_1_1_1,00.html)  
460 [html](http://www.oecd.org/document/40/0,3343,en_2649_34377_37051368_1_1_1_1,00.html)].
- 461 (28) EPA (1998). Health Effects Test Guidelines: OPPTS 870.1200 - Acute Dermal  
462 Toxicity. EPA 712-C-98-192. Washington, DC: U.S. Environmental Protection  
463 Agency. Available at:  
464 [[http://www.epa.gov/opptsfrs/publications/OPPTS\\_Harmonized/870\\_Health\\_Effec](http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-1200.pdf)  
465 [ts\\_Test\\_Guidelines/Series/870-1200.pdf](http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-1200.pdf)].
- 466 (29) ICCVAM (2009). Report on the ICCVAM-NICEATM/ECVAM/JaCVAM  
467 Scientific Workshop on Acute Chemical Safety Testing: Advancing In Vitro  
468 Approaches and Humane Endpoints for Systemic Toxicity Evaluations. Research  
469 Triangle Park, NC: National Institute of Environmental Health Sciences. Available  
470 at: [[http://iccvam.niehs.nih.gov/methods/acutetox/Tox\\_workshop.htm](http://iccvam.niehs.nih.gov/methods/acutetox/Tox_workshop.htm)].

471

## ANNEX 1

### PERFORMANCE STANDARDS FOR ASSESSMENT OF PROPOSED SIMILAR OR MODIFIED LOCAL LYMPH NODE ASSAY TEST METHODS FOR SKIN SENSITISATION

#### INTRODUCTION

1. The purpose of Performance Standards (PS) is to communicate the basis by which new test methods, both proprietary (i.e., copyrighted, trademarked, registered) and non-proprietary can be determined to have sufficient accuracy and reliability for specific testing purposes. These PS, based on validated and accepted test methods, can be used to evaluate the reliability and accuracy of other analogous test methods (colloquially referred to as “me-too” tests) that are based on similar scientific principles and measure or predict the same biological or toxic effect (11).

2. Prior to adoption of modified test methods (i.e., proposed potential improvements to an approved test method), there should be an evaluation to determine the effect of the proposed changes on the test’s performance and the extent to which such changes affect the information available for the other components of the validation process. Depending on the number and nature of the proposed changes, the generated data and supporting documentation for those changes, they should either be subjected to the same validation process as described for a new test, or, if appropriate, to a limited assessment of reliability and relevance using established PS (11).

3. Similar (me-too) or modified test methods proposed for use under this Test Guideline should be evaluated to determine their reliability and accuracy using chemicals representing the full range of the LLNA scores.

4. These PS are based on the ICCVAM/ECVAM/JaCVAM harmonized PS (9), for evaluating the validity of new or modified versions of the LLNA. The PS consist of essential test method components, recommended reference substances, and standards for accuracy and reliability that the proposed test method should meet or exceed.

#### I. Essential test method components

5. To ensure that a modified LLNA test method is functionally and mechanistically similar to the LLNA and measures the same biological effect, the following components should be included in the test method protocol:

1. The test substance should be applied topically to both ears of the mouse.
2. Lymphocyte proliferation should be measured in the lymph nodes draining from the site of test substance application.
3. Lymphocyte proliferation should be measured during the induction phase of skin sensitisation.
4. For test substances, the highest dose selected should be the maximum soluble concentration that does not induce systemic toxicity and/or excessive local

irritation in the mouse. For positive control substances, the highest dose selected should exceed the known EC3 values of the reference substances without producing systemic toxicity and/or excessive local irritation in the mouse.

5. A concurrent vehicle control should be included in each study and, where appropriate, a concurrent positive control should also be used.

6. A minimum of four animals per dose group is required.<sup>1</sup>

If any of these criteria are not met, then these performance standards cannot be used for validation of the modified test method.

## **II. Minimum list of reference substances**

6. The ICCVAM/ECVAM/JaCVAM harmonized PS (9) identified 18 minimum required reference substances and four optional reference substances (i.e., substances that produced either false positive or false negative results in the LLNA, when compared to human and guinea pig results, and therefore provide the opportunity to demonstrate equal to or better performance than the LLNA) that are included in the LLNA performance standards. The selection criteria for identifying these substances were:

- The list of reference substances represented the types of substances typically tested for skin sensitisation potential and the range of responses that the LLNA is capable of measuring or predicting;
- The substances had well-defined chemical structures;
- LLNA data from guinea pig tests and (where possible) data from humans were available for each substance; and
- The substances were readily available from a commercial source.

The recommended reference substances are listed in Table 1. Studies using the proposed reference substances should be evaluated in the vehicle with which they are listed in Table 1. In situations where a listed substance may not be available, other substances that meet the selection criteria mentioned may be used, with adequate justification.

---

<sup>1</sup> The ICCVAM/ECVAM/JaCVAM harmonized PS for the LLNA (9), which were developed in order to harmonise with the procedures described in OECD TG 429 to ensure international applicability, indicate that either pooled or individual animal data may be collected. However, because the proposed updated OECD TG 429 specifies the need for collecting individual animal data from a minimum of four animals per dose group, the provision to pool animals has been removed from this document.

536 **TABLE 1. RECOMMENDED REFERENCE SUBSTANCES FOR THE LLNA PERFORMANCE STANDARDS**

Number	Substance	CASRN	Form	Veh	EC3 (%) <sup>1</sup>	N <sup>2</sup>	0.5x - 2.0x EC3	Actual EC3 Range	LLNA vs. GP	LLNA vs. Human
1	5-Chloro-2-methyl-4-isothiazolin-3-one	26172-55-4	Liq	DMF	0.009	1	0.0045-0.018	NC	+/+	+/+
2	DNCB	97-00-7	Sol	AOO	0.049	15	0.025-0.099	0.02-0.094	+/+	+/+
3	4-Phenylenediamine	106-50-3	Sol	AOO	0.11	6	0.055-0.22	0.07-0.16	+/+	+/+
4	Cobalt chloride	7646-79-9	Sol	DMSO	0.6	2	0.3-1.2	0.4-0.8	+/+	+/+
5	Isoeugenol	97-54-1	Liq	AOO	1.5	47	0.77-3.1	0.5-3.3	+/+	+/+
6	2-Mercaptobenzothiazole	149-30-4	Sol	DMF	1.7	1	0.85-3.4	NC	+/+	+/+
7	Citral	5392-40-5	Liq	AOO	9.2	6	4.6-18.3	5.1-13	+/+	+/+
8	HCA	101-86-0	Liq	AOO	9.7	21	4.8-19.5	4.4-14.7	+/+	+/+
9	Eugenol	97-53-0	Liq	AOO	10.1	11	5.05-20.2	4.9-15	+/+	+/+
10	Phenyl benzoate	93-99-2	Sol	AOO	13.6	3	6.8-27.2	1.2-20	+/+	+/+
11	Cinnamic alcohol	104-54-1	Sol	AOO	21	1	10.5-42	NC	+/+	+/+
12	Imidazolidinyl urea	39236-46-9	Sol	DMF	24	1	12-48	NC	+/+	+/+
13	Methyl methacrylate	80-62-6	Liq	AOO	90	1	45-100	NC	+/+	+/+
14	Chlorobenzene	108-90-7	Liq	AOO	NA	1	NA	NA	-/-	-/*
15	Isopropanol	67-63-0	Liq	AOO	NA	1	NA	NA	-/-	-/+
16	Lactic acid	50-21-5	Liq	DMSO	NA	1	NA	NA	-/-	-/*
17	Methyl salicylate	119-36-8	Liq	AOO	NA	9	NA	NA	-/-	-/-
18	Salicylic acid	69-72-7	Sol	AOO	NA	1	NA	NA	-/-	-/-

Number	Substance	CASRN	Form	Veh	EC3 (%) <sup>1</sup>	N <sup>2</sup>	0.5x - 2.0x EC3	Actual EC3 Range	LLNA vs. GP	LLNA vs. Human
<b>Optional Substances to Demonstrate Improved Performance Relative to the LLNA</b>										
19	Sodium lauryl sulfate	151-21-3	Sol	DMF	8.1	5	4.05-16.2	1.5-17.1	+/-	+/-
20	Ethylene glycol dimethacrylate	97-90-5	Liq	MEK	28	1	14-56	NC	+/-	+/+
21	Xylene	1330-20-7	Liq	AOO	95.8	1	47.9-100	NC	+/**	+/-
22	Nickel chloride	7718-54-9	Sol	DMSO	NA	2	NA	NA	-/+	-/+

Abbreviations: AOO = acetone: olive oil (4:1); CASRN = Chemical Abstracts Service Registry Number; DMF = *N,N*-dimethylformamide; DMSO = dimethyl sulfoxide; DNCB = 2,4-dinitrochlorobenzene; EC3 = estimated concentration needed to produce a stimulation index of 3; GP = guinea pig test result (TG 406) (10); HCA = hexyl cinnamic aldehyde; Liq = liquid; LLNA = murine local lymph node assay result (TG 429) (1); MEK = methyl ethyl ketone; NA = not applicable since stimulation index <3; NC = not calculated since data was obtained from a single study; Sol = solid; Veh = test vehicle.

<sup>1</sup> Mean value where more than one EC3 value was available

<sup>2</sup> Number of LLNA studies from which data were obtained

\* = Presumed to be a nonsensitizer in humans based on the fact that no clinical patch test results were located, it is not included as a patch test kit allergen, and no case reports of human sensitisation were located.

\*\* = GP data not available.

**III. Defined reliability and accuracy standards**

7. The accuracy of a modified LLNA test method should meet or exceed that of the LLNA PS when it is evaluated using the 18 minimum required reference substances. The new or modified test method should result in the correct classification based on a “yes/no” decision. However, the new or modified test method might not correctly classify all of the minimum required reference substances. If, for example, one of the weak sensitizers were misclassified, a rationale for the misclassification and appropriate additional data (e.g., test results that provide correct classifications for other substances with physical, chemical, and sensitizing properties similar to those of the misclassified reference substance) could be considered to demonstrate equivalent performance. Under such circumstances, the validation status of the new or modified LLNA test method would be evaluated on a case-by-case basis.

**Intra-laboratory reproducibility**

8. To determine intra-laboratory reproducibility, a new or modified LLNA test method should be assessed using a sensitizing substance that is well characterised in the LLNA. Therefore, the LLNA PS is based on the variability of results from repeated tests of hexyl cinnamic aldehyde (HCA). To assess intra-laboratory reliability, threshold estimated concentration (EC<sub>t</sub>) values for HCA should be derived on four separate occasions with at least one week between tests. Acceptable intra-laboratory reproducibility is indicated by a laboratory’s ability to obtain, in each HCA test, EC<sub>t</sub> values between 5% and 20%, which represents the range of 0.5-2.0 times the mean EC<sub>3</sub> specified for HCA (10%) in the LLNA.

**Inter-laboratory reproducibility**

9. Inter-laboratory reproducibility of a new or modified LLNA test method should be assessed using two sensitizing substances that are well characterised in the LLNA. The LLNA PS is based on the variability of results from tests of HCA and 2,4-dinitrochlorobenzene (DNCB) in different laboratories. EC<sub>t</sub> values should be derived independently from a single study conducted in at least three separate laboratories. To demonstrate acceptable inter-laboratory reproducibility, each laboratory should obtain EC<sub>t</sub> values of 5% to 20% for HCA and 0.025% to 0.1% for DNCB, which represents the range of 0.5-2.0 times the mean EC<sub>3</sub> concentrations specified for HCA (10%) and DNCB (0.05%), respectively, in the LLNA.

577

## ANNEX 2

578 **Accuracy:** The closeness of agreement between test method results and accepted reference  
579 values. It is a measure of test method performance and one aspect of relevance. The term is  
580 often used interchangeably with “concordance” to mean the proportion of correct outcomes  
581 of a test method.

582 **False negative:** A substance incorrectly identified as negative or non-active by a test method,  
583 when in fact it is positive or active.

584 **False positive:** A substance incorrectly identified as positive or active by a test, when in fact  
585 it is negative or non-active.

586 **Hazard:** The potential for an adverse health or ecological effect. The adverse effect is  
587 manifested only if there is an exposure of sufficient level.

588 **Inter-laboratory reproducibility:** A measure of the extent to which different qualified  
589 laboratories, using the same protocol and testing the same substances, can produce  
590 qualitatively and quantitatively similar results. Inter-laboratory reproducibility is determined  
591 during the prevalidation and validation processes, and indicates the extent to which a test can  
592 be successfully transferred between laboratories, also referred to as between-laboratory  
593 reproducibility.

594 **Intra-laboratory reproducibility:** A determination of the extent that qualified people within  
595 the same laboratory can successfully replicate results using a specific protocol at different  
596 times. Also referred to as within-laboratory reproducibility.

597 **Me-too test:** A colloquial expression for a test method that is functionally and  
598 mechanistically similar to a validated and accepted reference test method. Such a test method  
599 would be a candidate for catch-up validation. Interchangeably used with similar test method.

600 **Performance standards (PS):** Standards, based on a validated test method, that provide a  
601 basis for evaluating the comparability of a proposed test method that is functionally and  
602 mechanistically similar. Included are; (i) essential test method components; (ii) a minimum  
603 list of Reference Chemicals selected from among the chemicals used to demonstrate the  
604 acceptable performance of the validated test method; and (iii) the similar levels of accuracy  
605 and reliability, based on what was obtained for the validated test method, that the proposed  
606 test method should demonstrate when evaluated using the minimum list of Reference  
607 Chemicals.

608 **Proprietary test method:** A test method for which manufacture and distribution is restricted  
609 by patents, copyrights, trademarks, etc.

610 **Quality assurance:** A management process by which adherence to laboratory testing  
611 standards, requirements, and record keeping procedures, and the accuracy of data transfer,  
612 are assessed by individuals who are independent from those performing the testing.

**Reference chemicals:** Chemicals selected for use in the validation process, for which responses in the *in vitro* or *in vivo* reference test system or the species of interest are already known. These chemicals should be representative of the classes of chemicals for which the test method is expected to be used, and should represent the full range of responses that may be expected from the chemicals for which it may be used, from strong, to weak, to negative. Different sets of reference chemicals may be required for the different stages of the validation process, and for different test methods and test uses.

**Relevance:** Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method.

**Reliability:** Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility.

**Skin sensitisation:** An immunological process that results when a susceptible individual is exposed topically to an inducing chemical allergen, which provokes a cutaneous immune response that can lead to the development of contact sensitisation.

**Stimulation Index (SI):** A value calculated to assess the skin sensitisation potential of a test substance that is the ratio of the proliferation in treated groups to that in the concurrent vehicle control group.

**Estimated concentration threshold (EC<sub>t</sub>):** Estimated concentration of a substance needed to produce a stimulation index that is indicative of a positive response.

**Estimated concentration three (EC<sub>3</sub>):** Estimated concentration of a substance needed to produce a stimulation index of three.

**Validated test method:** A test method for which validation studies have been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose.